REPORT OF CONCLUSIONS OF AEPD-ENISA'S EVENT ON DATA SPACES

DATA SPACES IN EU: Synergies between data protection and data spaces, EU challenges and experiences of Spain

I. INTRODUCTION

This report aims at presenting an excerpt of the topics discussed at the event: "*DATA SPACES IN EU: Synergies between data protection and data spaces, EU challenges and experiences of Spain*". The event was jointly organized by the Spanish Agency for Data Protection (AEPD) and the European Agency for Cybersecurity (ENISA) where related conclusions mainly derived from specific discussion sessions organized during the conference. This event was organised and conceived as a mechanism to make an analysis of the new EU regulation in digital and data domains and its interplay with data protection. We understand this conference as a useful exercise allowing a better implementation of the new regulation based on the concrete case of an EU Member State.

Since early announcement of the event, the response from numerous stakeholders was positive acknowledging the need to facilitate targeted discussions concerning implementation of EU regulation on data spaces. The community of stakeholders who attended this event represented a varied typology of sectors and topics and discussions and conclusions reflect such multidisciplinary. However, due to the complexity of the new ecosystem of data-access, the work to come will require a bigger effort targeting more ambitious and inclusive levels of engagement. In any case, this event allowed to demonstrate the added value of multidisciplinary interactions.

Data spaces regulation will cover a substantial part of technologies. Complementary, the GDPR highlights the need for a more solid and coherent framework for data protection in the European Union, backed by strict enforcement. In this regard, GDPR shall not been considered as a minimum or formal compliance requisite but as a mechanism to protect the fundamental rights in an effective way allowing the control by the citizens of their own data and to generate a trust in a dynamic internal market affecting all economic sectors.

It is evident that each one of these technologies will be addressed in data spaces. In other words, we could say that data spaces are a common hub where all technologies will be converge and on which new benefits for our society will be possible, provided that there is an adequate trust environment. However, it is difficult to have deep knowledge of the implications derived from the use of emerging technologies. Our responsibility as members of the data spaces is to know the impacts or threats that they could entail for our rule of law in general and for the rights and freedoms of each person or group of people in particular. In this sense, Control Authorities are aware of the need to equip themselves with experts who identify existing threats in each technology while proposing practical solutions and advice to guarantee the rights and freedoms of natural persons.

The development of data spaces must be associated, among others, with new advances in privacy (e.g. compute-to-data strategies, the federated processing, differential-privacy or the generation of synthetic data, etc.). Nowadays, there are many initiatives and new lines of work that may be synergised with data spaces. Data-space related actions should involve all sectors including research, industrial and academic. However, it is important to consider that when it comes to rights and freedoms of natural persons and their related data, personal data protection must be guaranteed.

In the European Union, the importance of providing trust to data spaces must be emphasised: trust that can be achieved by understanding the data protection regulations as a work tool to guarantee trust from the design with the necessary transparency and respect for the ethical and social values of our rule of law.

There are many data-oriented challenges ahead that will have to be addressed in the context of data spaces and related regulation. Nonetheless, the opportunities are infinitely greater than the issues and multidisciplinary interaction represents and added value for implementation. If something should characterize data spaces, it should be the collaboration between the stakeholders' ecosystem of a specific data space where different views can be conjugated around a common objective. With this enthusiasm, both AEPD and ENISA wanted to organise this event, aiming at providing useful considerations for the way forward.

II. MAIN PANELS

This section contains the main ideas that each of the panellists participating in the event wanted to convey through their participation.

A. WHY DO WE SPEAK ABOUT GDPR IN DATA SPACES?

This first panel counted on two speakers representing two supervisory authorities in their similar roles as heads of the unit. They are in charge of analysing the impact of technology and innovation on data protection in each of these institutions, who highlight the importance of data protection in these mass data accessing scenarios.

1. Luis de Salvador Carrasco, AEPD

> *As asset, data has the same significance than any other entity asset. We should expect that an enterprise, regarding its data assets, will be willing to join to data-access sharing initiatives that keep under control its know-how, market share, intellectual property, business secrets, competitiveness and the compliance and ethical principles. That control will give the enterprise trust enough to be an actor into the data-access sharing market.*
>
> *That control, and the trust the stakeholders need in the data-access sharing economy, is called "data sovereignty". The data sovereignty of the enterprises, the*

researchers, the States (that manage assets/data that belong to the citizens) and the natural persons is the way of "creating the trust that will allow the digital economy to develop across the internal market".

The way to get an effective "data sovereignty" means to implement an infrastructure open and federated, based in governance, policies, rules and standards, that allows to generate trust in all stakeholders by an effective control of their data assets by means of management, legal and technical tools. This is called a Data Space. Data Spaces must allow access to data, considering that access means "data use, in accordance with specific technical, legal or organizational requirements, without necessarily implying the transmission or downloading of data" (Article 2(13) DGA). Data access doesn't mean data dissemination, and of course, it doesn't mean uncontrolled data leaking. Data access means to implement ways to extract information, useful for an intended context, from different data sources with the purpose of creating value.

Management and use of Privacy Enhancing Technologies (PETs) can fulfil additional purposes beyond data protection. PETs can also fulfil several requirements of governance in a Data Space and work like "dual use" tools: GDPR requirements and other requirements that derive from the concerns of enterprise, public bodies, EU market sustainability, EU research and State security. The integration of Privacy tools and PETs in the governance model should be done by design of the Data Spaces. Such a way, they can work like "dual use" tools that facilitate the implementation of data sovereignty and the trust of the stakeholder to join the data-access sharing. Therefore, DPOs with a deep knowledge about data management and privacy by design tools must be involved in the design of Data Spaces to catch up with what is laid down in GDPR: control of the own data, trust in data-driven economy, legal certainty for all stakeholders.

2. Luis Velasco, EDPS

The GDPR stands as a cornerstone in shaping the data-driven future of Europe, not as an impediment but as a keystone in the burgeoning European Data Spaces. Data Economy stakeholders must find ways to merge economic growth with the protection of fundamental rights. As the data economy is poised to become a vital part of the EU's GDP and a significant employer, GDPR acts as a balancing force, ensuring that the pursuit of economic potential does not overshadow the necessity of upholding individual rights.

There is a widespread recognition among stakeholders—ranging from policymakers to industry leaders and citizens—that 'safe data realms' facilitated by GDPR are critical. These realms, or 'data spaces', are necessary for the data economy. They enable the smooth flow of data while simultaneously preserving privacy and other fundamental rights, setting a precedent for trust and safety in the digital age.

The European Data Protection Board (EDPB) and the European Data Protection Supervisor (EDPS) have been proactive in assessing legislative proposals from the Commission to shape the data economy. Their focus remains on protecting the core tenets of the GDPR, preventing any redefinition that could lead to legal ambiguity and the complexities that might emerge from the creation of new regulatory frameworks. They particularly emphasize the importance of clarity and stricter controls in sensitive areas, such as the secondary use of health data, to prevent infringements on individual rights and ensure alignment with GDPR standards.

B. EUROPEAN DATA SPACES INITIATIVES

The second panel, led by the European Data Protection Supervisor (EDPS), aimed to discuss European data spaces and the role of data protection and fundamental rights. Health data space was recurrently addressed not only because it has been the first EU regulated space, but also because health constitutes a data space in which processed data is particularly sensitive, e.g. collected as special categories of data under the GDPR.

1. Xabier Lareo López de Vergara, EDPS

The European Health Data Space (EHDS) it is the pioneer of the 10 data spaces foreseen by the Commission in its data strategy. Furthermore, it also deals with data that are both sensitive and labelled by the GDPR as special categories of data. Focusing on the EHDS was an easy decision.

The panel provided a view on the EHDS from three different angles.

First, Owe Langfeldt (European Commission's DG SANTE) drove us through the legislative proposal, currently under discussion by the European Parliament, the Council and the Commission. Owe explained the structure and main provisions of the EHDS Regulation proposal. A proposal that aims at improving the access to and exchange of electronic health data for the provision of healthcare services and for research & innovation and regulatory purposes.

Then, Jan Penfrat from European Digital Rights (EDRi) voiced EDRi's concerns about the legislative proposal and highlighted what they consider its main issues: insufficient user control, a too broad definition of health data, and unclear permitted purposes for processing.

Finally, Carlos Parra Calderón (Institute of Biomedicine of Seville) gave an overview on IMPaCT and HelathyCloud, two projects that explore how to construct IT infrastructure to achieve effective and secure health data sharing across Europe.

2. Owe Langfeldt, DG SANTE – EC

Owe Langfeldt, as expert in privacy issues related with the proposal of the new regulation of the European Health Data Space, talks about the key points of the regulation regarding personal data protection, as well as he makes an overview of other European data spaces.

The European Health Data Space is conceived for improving the health outcomes for patients and costs saving in the health system for Administrations. It is expected to be published by the end of the Spring of 2024.

He has encountered difficulties for data subjects in accessing their data and, also, similar difficulties for the health professionals. Besides, the interoperability between different national health systems is still improving with implementing the product legislation for electronic health record systems and the necessary technology. He notes that so far there are 11 countries active in ensuring interoperability and data exchange between member states.

He addresses the importance of using anonymised or at least pseudonymised data for research and the relevant role of the administrative body that will manage access to data for research. He also makes specific mention of market entry barriers due to the lack of competition of system providers.

Finally, he makes a review of the current and heterogeneous European Data Spaces projects, from different sectors, emphasising on the importance of the European Health Data Space in term of privacy due to the kind of data categories.

3. Carlos Parra Calderón, Institute of Biomedicine of Seville

The speech "European Data Spaces Initiatives: working on Trustworthy Health Research Data Infrastructures for the Success of Data Spaces" presented an overview of the security and data protection aspects to be taken into account in data spaces for health and biomedical research based on the experience gained in two initiatives, one national and the other European, taking into account the current regulations, the responsibilities of the treatment and the risks involved, all of which must be foreseen in the design of the infrastructures and taking into account as a critical aspect the profound needs of trust of health data providers with these infrastructures.

The national initiative is the Data Science program of the Precision Medicine research infrastructure associated with Science and Technology "IMPaCT," promoted by the Carlos III Spanish National Institute of Health, which defines a set of recommendations for handling sensitive data based on the Spanish National Security Scheme. The European initiative is the coordination and support action funded by the European Commission "HealthyCloud," which defines a strategic agenda for the Health Research and Innovation Cloud in Europe and defines a series of services to support the applicable legal and regulatory framework.

4. Jan Penfrat, European Digital Rights

Jan Penfrat, as the representant of the civil society, talks about data spaces as a fundamental shift from the principles of GDPR, and in particular, the Health Data Space because it concerns very sensitive personal data. He prefers talking about how making GDPR and data spaces compatible, instead of speaking about synergies.

He does not agree with the statement that data is the new oil, at least in the way it has been presented, although he does agree that personal data is as toxic as oil can be, so we will have to collect as little as we can (data minimisation and purpose limitation). In this respect, he considers that the European Commission's initial position in its proposal has been changed. We are talking now about data as an asset for companies, data market and strategic assets as a result of the market created by the DMA, despite the fact that this regulation states that the transfer of data between gatekeepers and non-gatekeepers should not contain personal data.

All of these scenarios are based on a vast amount of data gathered on a voluntary basis. However, it has been identified that a vast majority of European citizens would not be comfortable with the sharing of their medical data. EHDS is supposed to give more control to the patient. However secondary use can in some ways do the opposite and give control of your data to third parties who define themselves as having a research interest. On the one hand, as citizens we expect secrecy with doctors, but on the other hand, for secondary use it is completely out of the control of the data subject who and for how long their data is used. Innovation can not out rule fundamental rights.

> *Personal information is not a commodity, it is a representation of our right to privacy and commodifying it bring us a rabbit hole that I think very quickly will make our fundamental right to privacy obsolete.*

He highlights as a conclusion the proper management of the patient's informed consent and the opt out option.

C. INTERPLAY GDPR-DGA-DA-DMA-DSA-EHDS-AIA IN DATA SPACES

The third panel was led by the EDPB and addressed the impact of the new digital package regulation on data spaces and data protection from an EU perspective. All these regulations will have a very close relationship in these scenarios of accessing to massive data, the data spaces. Accordingly, this panel counted on three European experts on data protection, with strong background on analysing regulation and technology.

1. Anna Lytra, EDPB

Anna Lytra, from the EDPB data protection officer's office team, wants to address some aspects of the new regulations of the European digital package that have a place in the data spaces with this panel. To this end, she has three distinguished speakers in the field of privacy applied to digital technologies.

2. Marit Hansen, Schleswig-Holstein Data Protection Authority

The intervention of Marit Hansen, recognised for her background in privacy and data protection with a technical profile as well as extensive legal knowledge from her long professional career, focuses on pointing out several high-level aspects that have to do with the new digital regulatory package and the GDPR.

Only one sentence: with the new European Acts, GDPR remains unaffected.

Developments and applications resulting from the new regulations will have to be adapted to comply with the GDPR, so workable solutions will need to be developed.

A primary aspect will be to address the terms of risk detection and risk mitigation when analysing the interplay between GPDR and the new Acts, particularly the AI Act, which will influence on how to shape technology and organizations to the risks of rights and freedoms of natural persons. In this regard, a “professionalized” fundamental rights impact assessment (FRIA) should be carried out.

Finally, as a final remark, as most of new Acts create their own supervisory authority with its own legal terminology, the orchestration between them will be a key issue.

3. Regina Becker, Luxembourg National Data Service

The creation of data spaces is one of the main goals of the EU Strategy for Data. Recent legislative acts are the Data Governance Act and two draft Regulations, the Data Act and the European Health Data Space. However, when it comes to creating data spaces for secondary use, a space with harmonised data available under a harmonised data governance, it becomes apparent that these Regulations do not provide a legal basis to harmonise and hold data for secondary use.

Alternative ways to create harmonised data spaces for secondary use of sensitive personal data are not easy to implement. Most entities that have collected data for their primary purposes neither have a legal basis to harmonise data for secondary use nor to share them systematically for users’ purposes. Their mission and thus legal basis is entirely focussed on their own primary tasks. Even where entities have a mission to make data available for secondary use, each entity has its own data governance by law, which leads to a fragmentation across Europe. Consent with its requirement to be informed and specific is not suitable for the disclosure to users either.

Harmonised data spaces for sensitive personal data need European law. A potential solution is offered through the new legal instrument of a European Digital Infrastructure Consortium (EDIC) introduced in the Digital Decade Policy Programme (DDPP). EDICs have legal personality and are created through a Commission implementing decision. Where an EDIC becomes the controller for the data disclosure, the implementing act should provide the legal basis based on the EDIC’s mission. However, the legal framework of the EDIC as defined in the DDPP leads to questions with respect to the sufficiency of the implementing act, a setback that still needs solving.

4. Ricard Martínez Martínez, University of Valencia

Ricard Martínez Martínez, as an active privacy expert in the field of technology, starts his keynote speech with the following thoughts after having heard other previous panellists:

> *This is not just about data protection, we are talking about the future of data driving public policies, data driving welfare state, data driving healthcare services, data driving society. According to GDPR, data is addressed to promote human being, to promote common good, and this will be our approach. This is not about forbidding processing personal data, this is about processing data in a secure environment with legal warranties.*

Regulations such us DGA, EDHS or DA consist of empowering citizens, empowering data subjects, which is not an easy task because: 1) data subjects don't understand privacy policies, 2) we are speaking about services that are, in practice, form monopoly, and 3) there is an unbalanced situation, particularly in internet services, or the real network of health care research.

Having established the boundary conditions, it is considered of vital importance to work on the following aspects: dynamic consent (different from the traditional way to give consent or to control data), free and unambiguous consent in the digital world, management of reputational risks (which can improve trust to the society), dissemination of what we are doing, deal with an ecosystem where consent is not the most suitable legal basis (we have to bring the idea of common good back), working from a public interest perspective (not from an individualistic one), fostering the use of privacy-enhancing technologies, warranting the secure, traceable and available local infrastructures to ensure that all local nodes are able to work together in a federate way, implementation of legal governance contractual clauses between the different stakeholders involved at all levels (including terms related to ethics), and the improvement of the supporting staff.

D. DATA SPACES AT A NATIONAL LEVEL THROUGH EUROPE

The fourth panel, led by the Spanish Data Office, wanted to address the national data spaces that are already being developed. For this reason, the European Research Centre of the European Commission presented his role as supporting institution for the development of data spaces and two examples of Spanish data spaces.

1. Alberto Palomo, Secretariat of State of Digitalisation and Artificial Intelligence

> *The purpose of the panel "Data Spaces at a national level through Europe" was to show practical examples of data spaces in Spain, as well as the general context for their deployment and impact as per the analysis and recommendations made by the Joint Research Centre. Moderated by Alberto Palomo (Spanish National Data Office), the panel started with Eimear Farrel (JRC) addressing the role that this unit within the European Commission has played in providing techno-socio-economic*

perspectives around data-sharing, alongside non-binding recommendations and good practices. They have also produced various resources for data space requirements, including a map of resources and an open repository of knowledge, that provides a holistic view for data space stakeholders. Following, Alberto shared the view from Gaia-X, a uniquely positioned cloud & data initiative that has various lighthouse projects already in-flight, and shared the importance of an adequate multi-level governance framework across projects, industries, Member States and EU-wide level.

At last, Rocío Báguena and Maite Ambrós —respectively from the Spanish Ministry of Transport (MITMA) and the Spanish Ministry of Agriculture (MAPA)— presented a view of the strategic role that data spaces play across their units, as well as projects midcourse or already operational. Both ministries are eager to promote a sustainable production model for data-driven innovation. In summary, these dialogues evidenced that efforts are being conducted towards the effective deployment of data spaces across different sectors, whereby European organizations are digitalizing their value chains thanks to reliable and traceable data, thus generating a competitive advantage across international markets.

2. Eimear Farrell, EC Joint Research Centre

Eimear Farrell, as a scientific expert in the field of data, focuses her exposition on the work developed by the Joint Research Centre as the science and knowledge centre of the European Commission.

She is involved in the development of technology which make it possible to put together both the policies and the implementation of the DGA, DA, DMA, AIA and all the digital regulations. The JRC's catalogue has more than 3000 datasets for the research, and they have published more than 500 documents related with data sharing.

The last publication where she has actively contributed is related with data spaces: "European Data Spaces - Scientific Insights into Data Sharing and Utilisation at Scale"[1]. She remarks the mapping of the landscape of intermediaries which play an important role in the DGA and in the data spaces and a list of functional and non-functional requirements of data spaces, the way they operate and their quality attributes. She also points out that they are also working on a new publication about PETs.

She is also involved in supporting the European Commission in the design and deployment of the Green Deal data space, where the Gaia-x building blocks are being tested to check compatibilities in this environment.

3. Rocío Báguena Rodríguez, Spanish Ministry of Transport

Rocío Báguena, as expert in transport technology and data, talks about the role of the Ministry of Transport and Urban Agenda regarding data and data sharing and refers it to

[1] https://publications.jrc.ec.europa.eu/repository/handle/JRC129900

the "National strategy about the safe and secure, sustainable and connected mobility 2030". The strategy includes a total of 8 pillars and up to 150 different measures, and 2 of these 8 pillars are fundamental to data and data sharing: pillar 5 related to smart mobility (or transport of a person) and pillar 6 smart intermodal logistic chains. They include several measures and linked projects.

She also speaks about the national access point for the multimodal transport which gathers information from both private and public companies. It has contributed to the open data system with over 100 dataset that can be used for research and statistics.

4. Maite Ambrós, Spanish Ministry of Agriculture, Fisheries and Food

The uptake of digitalization in the agrifood sector is increasing, data sharing initiatives are widespreading, nevertheless the sector is still not familiar with the concept of data space. The Common Agricultural Policy has focused too much on monitoring and control of the public funds and data to do so (secondary use), whereas the primary use of data has to be further enhanced, to improve profitability, environmental performance and help primary producers hold a stronger position in the asymmetric agrifood chain.

There are a few promising initiatives that could turn into real data spaces, many of them building up around data cooperatives as data intermediaries that assure an environment of trust, shared goals and inclusive governance. One example is carried out by the "Spanish Confederation of AgriFood Cooperatives"[2] *that is capturing data needed to apply for CAP subsidies, in an agricultural holding digital register or "cuaderno de campo digital"*[3] *but with the idea to improve it with a GIS system, and to benchmark cooperatives agricultural practices against one another. In other Sanish examples breeders associations are the data cooperatives, that gather the data from the everyday practice and production of livestock farmers as well as from researchers, family trees on filiation, and artificial insemination centers like the project "GC4 Sheen"*[4] *where a Federated Data Cloud Platform with Artificial Intelligence Layer for the Genetic and Reproductive Improvement of the National Dairy Sheep will be developed.*

The Coordination and Support Action AgriDataSpace is mapping these initiatives in the whole EU and defining common building blocks of data spaces from multiple angles (technical/technological, business, and organisational/operational) but with the perspective of farmers, SMEs, and particularities of the Agrifood sector[5]*.*

[2] https://www.agro-alimentarias.coop/

[3] https://www.agro-alimentarias.coop/posts/cooperativas-agro-alimentarias-de-espana-e-hispatec-presentan-el-cuaderno-de-campo-cooperativo

[4] https://gc4sheep.com/

[5] https://agridataspace-csa.eu/

E. DATA PROTECTION BY DESIGN AND BY DEFAULT TECHNIQUES IN DATA SPACES

Finally, the fifth panel, led by ENISA, aims to provide a vision of how to take data protection by design and by default into account in these massive data accessing scenarios, by presenting the view of three experts in data protection from different areas of expertise.

1. Prokopios Drogkaris, ENISA

Data Governance Act creates a framework where Data Holders, Data Intermediaries, and Data Users cooperate to ensure the responsible and compliant sharing, processing, and use of data, including personal data. Data protection by design and by default are two pillar principles towards protecting individuals' rights and freedoms and meeting GDPR requirements.

However, is the practical deployment of these principles something completely new or we can draw lessons from the experience that we already have in the existing processing operations? To what extent can technical standards support us in that process and which are the more specific elements that we need to consider? What is the state of the art in technical and organizational measures that can support engineer data protection in EU Data Spaces?

2. Isabel Barberá, Rhite

Trust, Synergy, and Interoperability are common concepts in data spaces. While they are essential, they often focus on building efficient systems and not always on the critical synergy with users, end-users and data subjects.

For trust to flourish, the feeling of safety is paramount. PETs and rigorous data quality systems are valuable safeguard components, yet they alone cannot guarantee trust. To truly establish trust we need transparency. This can only be achieved by recognizing the value of data and the common interests of all participants. This shared understanding is what can create synergy and interoperability.

Exploring the data spaces threat landscape leads to some concerns. While data intermediaries have the potential to foster trust and synergy, who should select them? Participants, member states, broader European entities? Embedding data subject rights in data space architectures and addressing operational costs disadvantages for SMEs are also critical concerns. Interoperability relies on standards, yet governance remains an open question, who is accountable? Who is responsible for identifying and mitigating risks?

Perhaps it's time to add an "E" for "Ethical" to the FAIR acronym (Findable, Accessible, Interoperable, Reusable) to underline the significance of ethics and the protection of fundamental rights in the data spaces discourse.

3. Marie Charlotte Roques Bonnet, Data protection legal advisor

The creation of a consistent technical and organizational framework supporting efficient sharing of personal data within and across all relevant sectors in EU (EU common data spaces) is essential for those to be efficient. All data holders willing to promote re-use of personal data for social and economic good, whether they are public or private entities, controllers or processors, must demonstrate accountability, by means of, as applicable, revamped internal mechanisms, data sharing agreements and sensible privacy management programs (PMPs). To do so, some building blocks of accountability are:

1. *Clear-cut identification of responsibilities and obligations for data holders and users.*
2. *Effective internal governance of personal data sharing.*
3. *Cooperative external governance of personal data sharing*
4. *Addition of a dedicated "Data sharing program" section in the PMPs of data holders.*
5. *Design of targeted Data Sharing Accountability tools to reduce the risks (i.e. "data altruism mechanisms").*
6. *Balanced security / risk-mitigation objectives to ensure a sufficient quality of data to be shared.*
7. *Ethical assessment of envisaged data sharing practices, for both data holders and data users.*
8. *Transparent information sharing between data holders and data receivers.*
9. *Contractual framing of data sharing practices.*
10. *Transparency towards individuals.*

4. Irene Kamara, Tilburg Law School

To achieve data protection by design and by default in European data spaces standardisation has an important role to play. Technical standards in general foster trust among different actors in data spaces and provide common baseline benchmarks. The Data Governance Act 2022/868 in its Recital 23 already points towards technical standards, codes of conduct, and certification as good practices.

In the field of data protection and privacy by design and by default, European and international standardisation organisations have published several standards that are of relevance also to data spaces. In specific, several requirements, controls, and processes of existing data protection by design and by default technical standards are relevant both for the governance/organisational layer and technical layer of data spaces. Those include consumer/data subject *communication requirements, risk assessment methodologies, data breach response plans, system and architecture requirements definition.*

There are however also some challenges in using existing technical standards. Those include for example that current technical standards on data protection by design and by default focus on organisations, rather than the processing lifecycle across organisations. Furthermore, specific approaches and techniques relevant to data spaces such as the compute-to-data approach, which is provided also in the European Health Data Spaces proposal, are not reflected in existing data protection standards.

III. DISCUSSION TABLES

The final part of the event focused on a series of working tables in which attendees discussed issues related to data protection in data spaces.

These tables were open for discussion, to raise open questions, identify the main problems as well as to propose solutions or, at least, identify where further work would be done to reach them.

Each working table was led by two moderators who were in charge of chairing the discussions, compiling a summary of outcomes and presenting in the plenary session that concluded the event.

The eight working tables that were set up addressed the following topics. Each table counted on a varied representation of 10 people on average:

1. Actors, Stakeholders and Roles
2. Risk management and DPIA
3. Data Subjects’ Rights
4. Transparency and Accountability
5. Data Protection Officer
6. Enforcement and Supervisory Authorities
7. Technologies for data sharing
8. Data Breaches and security measures

A. ACTORS, STAKEHOLDERS AND ROLES

Ricard Martínez Martínez (Universitat de València) and Jesus Rubí Navarrete (AEPD) as moderators of the “Actors, stakeholders and roles” working table reached to the following conclusions after debating with the participants of this working table.

The issues addressed in this working table highlight the main challenges of any project to define a data space and the use cases in it. The identification of the roles of each actor in a data space is key to establishing the appropriate governance mechanisms and identifying data protection responsibilities, among others.

Based on current experiences, it is not possible to define a priori the roles deployed in a Data Space. This is due to the complexity of its variety of activities and structure. A quick example of considerations that can begin to be listed for defining/designing a data space could be:

A. Ordinary Administrative Management.
   - Data Space User Management: registration of data access users, consultation of dataset catalogues or application dashboards, subscription to newsletters, etc.
   - Engagement of Data Holders or Federated Nodes: negotiation process for joining the data space.
   - Management of cookies, social media, communication channels, events, dissemination.
   - Internal management of human resources.

B. Data Space Management in the provision of processing services. In this environment, all possible relationships can occur, including:
   - Successive Controllers (recipients in a disclosure by transmission or by a simple data query: data holder to platform, data holder to data user disclosures or data sharing).
   - Joint-Controllership (consortium management of the data space by data holders, use of data by a research consortium or data access users and data holders).
   - Support for storage or processing by one of the nodes (data processors).
   - Anonymization scenarios: support for the anonymization process (data processor) and sharing and usage of anonymized data supported by specific legal agreements such as Data Sharing/ Transfer Agreements (Data Holders) and Terms & Conditions (Data Access Users).

Therefore, Data Spaces must necessarily have a role definition model based on the EDPB's Guidelines 07/2020 on the concepts of controller and processor in the GDPR.

Consideration should be given to the possibility of a single individual (natural or legal person) assuming multiple roles that could potentially influence decision-making processes. Therefore, the definition of roles should include governance processes for declaring and resolving any conflicts of interest.

As highlighted in the previous considerations, data space models, particularly federated environments, define relationships of high complexity. Furthermore, these models can evolve and undergo changes for various reasons. For example, a data space federation composed of a consortium of partners may transform into an independent legal entity. The addition of new data holders or the introduction of new processing services can also impact the roles deployed in a data space.

Moreover, the creation of bodies or entities by different regulations does not necessarily imply a clear definition of their roles. For instance, if a region with its own administration creates a data space for the exploitation of its information for primary or secondary purposes, it operates as a controller. If it also offers processing services to municipalities within its territory, it might act as a processor. However, if the purpose is to deploy data analysis for territorial policy design, there could be successive controllers or joint controllers. On the other hand, in an infrastructure open to all stakeholders, the definition of the purpose of secondary use falls to the data access user.

Unless in very clearly defined cases, the legislator lacks sufficient information for role definition. Additionally, if the design is flawed, it should result in the non-application of the law when it contradicts the GDPR. In this context, the risk of institutional conflicts or legal disputes should be avoided. Consequently, it is not advisable to define a priori roles by law.

In summary, the following statements are identified as key to take into account when defining/ designing a data space:

- The legal design of data space must contemplate, assume and implement the governance model that arises from GDPR.
- Each Data Space must implement procedures for a clear determination of GDPR roles and govern potential conflicts of interests.
- Legal predetermination of roles could be a risk. Joint controllership should be considered particularly risky due to the difficulty of its implementation.

Some additional reflections to be considered would be:

- Data users, data space and data holders should be particularly accountable in the definition of lawful basis for processing and applying GDPR compliance procedures. In particular, cooperation and transparency between all the actors on activities such as the Data Protection Impact Assessment or Artificial Intelligence risks impact assessment should be extremely relevant.
- Data Spaces must provide a clear institutional information about their procedures and governance model. The governance policy must be based in real evidence.
- Members of the working group consider that there are not adequate conditions of maturity for fulfilling GDPR at European Data Spaces.
- Enablers, technical staff are essential. Public bodies and future intermediation services should consider their roles, profiles, and job positions such as compulsory investment. The working group underlines that there are not enough human resources, internal private and public sector culture on GDPR compliance to provide certainty, security, and trust for society. This effort will be especially relevant for data altruism activities.

B. RISK MANAGEMENT AND DPIA

Isabel Barberá (Rhite) and Rafael Pastor Vargas (UNED) as moderators of the “Risk management and DPIA” working table reached to the following conclusions after debating with the participants of this working table.

These proposed open questions that drove the debate in the working table were: How to detect and manage risks created from pooling the data (e.g., linkage possibilities))? Who is responsible for mitigating those risks? Are there methodologies for assessing the re-identification risks associated with the data-sharing process considering the possible data types? What types of risk scenarios are common in data spaces? How should they be assessed and managed? Can general recommendations be given for DPIA, or are they specific for each data space according to the sector and data type processed? May quantitative risk assessment methods help to improve DPIA results?

The members of the working table considered that there is a complex relationship between the roles associated with data management/use. Furthermore, the greater the complexity of data processing, the greater the risk of data compromise. To mitigate these complexities, specific guidance is needed to deal with DPIAs. Having a complex environment requires specific DPIAs (per role/user, per data domain) and the global DPIA of the particular data space environment/ecosystem.

Considerations were also given to issue such as who should be in charge of risk assessment within a data space ecosystem. It could be the supervisor, the data intermediary or even the data processor, so again, having a system of decentralized risk assessment accountability seems necessary. It is necessary to distinguish which entity is responsible for the execution and supervision of the risk assessment, depending on the structure of the implemented data space. The standardization of DPIAs is another important issue, and it is considered a priority to have specific points that ask/explain aspects of risk analysis in the massive data-access sharing scenarios in a data space, which do not exist at present. In addition to this, there are considerations about the influence of fundamental rights on impact assessments (FRIA, Fundamental Rights Impact Assessment). This would increase quality and trust but it also requires more awareness, education, training and resources.

Another critical issue to consider is the risks of re-identification in data environments, where it is possible not to know precisely the semantic richness of the environment and to have the possibility of having information that allows this re-identification. It should be considered during the design/ideation process. Later, it could be tested with automated tools to facilitate efficiency and productivity in the data space. In addition, being a live environment with constant data updates, it is advisable to use dynamic evaluations and not traditional ones based only on time periods. Again, automation is a must.

As a summary of the contributions it can be concluded that it is necessary to follow the following lines for the future:

- Guidance to deal with DPIA's and its standardization is needed and a priority.
- Include FRIA (Fundamental Rights Impact Assessment) considerations in the risk evaluation.
- Provide (to have) a library of threats specific to data spaces, modelling complex relationships between data spaces' roles.
- Focus on assessment related to re-identification as a mandatory part of the DPIA.
- Complex data spaces need different DPIAs, so defining a DPIAs hierarchy and a clear definition of responsibilities in (dynamic) risk assessment in Data Spaces is essential.

C. Data Subjects’ Rights

Marie Charlotte Roques Bonnet (EU Privacy Legal expert) and Javier Gomez Prieto (ENISA) as moderators of the “Data Subjects Rights” working table reached to the following conclusions after debating with the participants of this working table.

In data spaces more than ever, the control of users over their data will determine how effective their data subject rights will be. Therefore, it is essential for them to know when their personal data: i) is processed for further purposes, ii) is processed in a de-identified or identifiable format. Based on this initial mapping, individuals shall be empowered to: 1/ share their reasonable expectations (see recitals 47 and 50 of GDPR), 2/ exercise their rights in practice, such as the right to easily consent or withdraw consent (that is to say in few clicks).

The group acknowledged that irrational handling of rights by data subjects would be an issue but admitted we should be enabled to exercise our rights-easily. To this respect it is fundamental to achieve a balance between providing personal data and getting a service (risk-based approach and impact-based approach). A proactive and accessible information shall be provided in a digestible way: build valid consent in a tangible way: explaining how they could exercise rights and avoid deceptive models / hidden activities. The group noted that public research (e.g. not profitable activity to companies) shall one of the main aims for data spaces.

In case personal data is processed in a directly identifiable format, appropriate Technical and Organisational Measures (TOMs) should be taken, and the obligations remain as demanding as in the GDPR. The example of security breaches impact, that might be higher in interoperable data spaces, raised the question of a right that could be specific to data spaces and consisting in exercising a right to have your personal data left out or never in. Another example was about Technical Decentralisation which is inherent to data spaces: storage should be sliced to divide and minimise risks.

All participants agreed on the decisive contribution of cybersecurity good practices and valued clear data spaces’ good practices (i.e. decryption keys stored by public bodies, access authorisation, interoperability specific risks, differential privacy). All participants were in favour of a strong pseudonymisation by-default. Anonymisation was assessed as likely to “not work”: singling out and inferring would be possible it is useless data because it is not qualitative enough to help research and innovative data processing.

Part of the discussion also addressed specific views related to the “EU Health Data Space Act”. One of the first observations was that having too many legislative tools would not, by nature and by definition, facilitate interoperability of such different data spaces, something which a priory goes against a smooth exercise of data subject rights, and not only on portability (article 20 GDPR). Such rights should be framed in practice, starting with a ground assessment and a bottom-up multi-sectors consultation phase. Key elements across these intersections would be addressed precisely through the notions of: a) “structured, commonly used and machine-readable format”, and b) where technically feasible”. Such operational brainstorming and screening could be driven using either targeted EU/international taskforces or through the review of sector-based good practices benchmarking.

Finally, contextual and operational constraints should be taken into account at all decision-making levels. On this front, the main takeaway of discussions could be summarised as follows: it is not a problem, practices would be different from a sector to another, and TOMs handled differently, but it is essential that rules and principles be an open standard easily replicable from a sector to another, from a data space to another. This approach will determine whether individuals are empowered in practice to exercise their rights in a simple and consistent way. In a nutshell, the legislative frameworks for Tech-friendly environments and data spaces should not specify data subject rights in a sector-based approach but just create consistent tools, from a sector to another, in order to enable them to simply exercise such rights.

D. TRANSPARENCY AND ACCOUNTABILITY

Javier Huerta Bravo (Cullen International) and Andrés Calvo Medina (AEPD) as moderators of the "Transparency and Accountability" working table reached to the following conclusions after debating with the participants of this working table.

Participants agreed that data spaces are quite nascent, in the early stages of development. Thus, at this stage it is difficult to frame all the accountability and transparency obligations in a data space context.

Participants also noted that, where relevant, accountability and transparency requirements under the EU General Data Protection Regulation (GDPR) might need to be complemented with accountability and transparency obligations set out in other relevant digital legislation, including the draft EU Artificial Intelligence Act. The increasing complexity of the EU regulatory framework on data was also addressed.

Discussions revolved around the concept of accountability and how this fundamental principle relating to the processing of personal data in the GDPR should be understood in the context of data spaces as an inherent part of them by design and some of the intervenient in data spaces should require assistance like it should be the case of startups and SMEs.

One of the participants pointed out that accountability means "going back to each data processing operation and being able to explain what was going on at a given time", to demonstrate compliance. However, participants emphasized that accountability is different from compliance, being rather an important aspect of the latter.

The above definition puts a strong focus on the traceability dimension of the accountability principle. Being able to trace back data processing operations becomes more relevant in the context of data spaces, as these are expected to involve many data processing operations, data controllers and processors, and data subjects.

Further, participants noted that there should not be a fixed or rigid accountability model for data spaces. Instead, accountability (or the applicable accountability models) should be dynamic, based on the state-of-the-art technologies, and flexible enough to be tailored to the specificities of each data space and of each personal data processing.

Different privacy enhancing technologies (PETs) and technical solutions adapted to the characteristics of each data space could be implemented. The proactiveness that is implicit in the accountability principle becomes crucial in this context.

Moreover, participants noted that, given that the line between personal and non-personal data is often quite blurred, the accountability rules in the GDPR should be applied to both personal and non-personal data.

Participants also addressed effective ways of ensuring transparency in data spaces. They agreed that transparency processes could be automated or semiautomated, and that researchers and the academia could assist organizations on this aspect.

Further, participants discussed how common standards, specifications, certification and labels could help bring transparency into data spaces. The European Commission's proposal for a health data space already flags some of these instruments and can inform other data spaces.

E. DATA PROTECTION OFFICER

Anna Lytra (EDPB) and Carlos Saiz (ISMS Forum) as moderators of the "Data Protection Officer" working table reached to the following conclusions after debating with the participants of this working table.

Several DPOs participated in this roundtable expressed that they face difficulties with identifying the role of the different actors involved in data spaces as controllers/processors/sub-processors. This can have an impact on the DPO's quality of advice to their controller/processor within the organisation on several matters.

The Spanish data protection law does not provide fines for the public administration in case of infringement. Considering this, the DPOs may face some difficulties on how to promote/encourage the compliance within their organisation.

The DPOs participated expressed that an EU network, where DPOs can exchange on their practices and challenges that they face in the data spaces context, would be much appreciated.

F. ENFORCEMENT AND SUPERVISORY AUTHORITIES

Enrique Factor Santoveña (AEPD) and Manuel González Seco (CTPD) as moderators of the "Enforcement and Supervisory Authorities" working table reached to the following conclusions after debating with the participants of this working table.

The main topics discussed in this working table were: governance, position of different and many authorities, innovation in data spaces and the need to deliver value.

Governance has been tackled by raising the following questing: "Should governance include hard enforcement or soft approach?" After a discussion among the participants, the following conclusions were reached:

- Soft approach has worked best, with bonus/malus system based on reputation. But there must be hard enforcement, as an ultima ratio.
- Importance of public-private collaboration, to apply enforcement/guidelines that are enforceable and effective.
- Need to extend of hard enforcement measures to public stakeholders: meaningful (non-financial) measures such as prohibition of processing that can be applied in data protection, which could have an important effect.
- SMEs should also be considered.

Going to the second topic discussed, related to the position of different and many Authorities, the participants concluded that:

- Coordination is necessary and it might be difficult when dealing with independent regulators.
- Overlapping spheres of competence. Interplay between different regulatory frameworks at different levels.
- Coordination between regulators: should not only coordinate but also drive collaboration and use of data for the right purposes. May be the case of receiving, for example, nine requirements from different authorities and at different levels (Regional, national, European).
- Reinforce talent acquisition: lessons learnt from cybersecurity that can be applied in this area.

Moving forward to the next topic “innovation in data spaces”, all participants agreed in stating that regulation is not a brake; it establishes a working environment that guarantees development while respecting fundamental rights. The national framework should not be a deterrent on the implementation of data spaces.

And, finally, when debating about the need to deliver value, the working table concluded that data spaces will only work if they deliver value. These new environments require powerful investment and expensive maintenance. The added value does not have to be economic, the final goal is to benefit all participants.

G. TECHNOLOGIES FOR DATA SHARING

Christina Michelakaki (FPF) and Miguel Peñalba Moldes (AEPD) as moderators of the “Technologies for data sharing” working table reached to the following conclusions after debating with the participants of this working table.

The working table shed light on critical issues and identified insightful conclusions. One central question that emerged was whether there are technologies capable of ensuring GDPR compliance when sharing data. Two distinct perspectives emerged: that of the industry and of the entities providing tools and technical means and that of the regulators. From an industry standpoint, it was discussed that companies are hesitant to share data using PETs or other technologies due to a lack of awareness about these solutions and due to lack of regulatory guidance on the matter. Many participants highlighted the fact that “success cases” in data sharing are not demonstrated enough resulting in businesses' unwillingness to trust new technologies. From a regulatory perspective, it was made clear that regulators are not in a place to indicate which technology works for data sharing purposes given that a measure that may be appropriate for a certain context could be ineffective for another one.

Another block of issues that emerged during the workshop concerned technical factors that go counter to data sharing in the specific context of a Data Space. First of all, there was a call for the development of tools and standards that ensure accountability and trust in data-sharing processes. Also, participants noted a lack of European technical solutions, observing that there are not many competitors in the EU area. The issue of data quality was also touched upon. Then, data maturity was deemed a secondary concern compared to the critical issue of interoperability. Problems arose when trying to use solutions provided by different PET providers due to a lack of interoperability. Thus, the development of standards was proposed as a means to instil confidence among all parties involved in data sharing.

The workshop moved on to explore specific examples and success stories in sharing data and the tools used for this purpose. In the context of health services, encryption in cloud-based systems was highlighted as a solution allowing computations to be performed directly on encrypted data without the need for decryption. However, the vast amount of data and complexity of contractual clauses posed considerable challenges. Then it was also shown that companies prefer the approach of a centralized data sharing where an instrumental company is obtaining consent from clients. However, it was agreed that this scenario comes along with inherent risks, such as the withdrawal of consent. Thus, it was suggested that a shift towards federated learning, where computation is brought to the data rather than moving data to a centralized location, could mitigate such risks. More specifically, federated learning enables model training on distributed data sources without sharing the raw data. This tool can be used by companies to work collaboratively without really sharing insights but getting advantages as well. Nonetheless, concerns about competition and potential monopolies persisted, highlighting the need to consider possible solutions not only from a GDPR compliance perspective but also from a competition law standpoint.

In conclusion, the participants shared their thoughts on the challenges, opportunities, and considerations surrounding data sharing and technologies. They emphasized the importance of awareness, standards development, and the need for versatile tools like federated learning to overcome challenges and enable safe and efficient data sharing in the future.

H. DATA BREACHES AND SECURITY MEASURES

Olga Rierola Forcada (APDCAT) and Irene Kamara (Tilburg Law School) as moderators of the “Data breaches and security measures” working table reached to the following conclusions after debating with the participants of this working table.

The high volume of personal and non-personal data processed in Data Spaces and the permanent interconnection between different systems for data sharing and exchange, increase the risk that personal data breaches occur (or increase the risk/probability of personal data breaches materialization). The participants of the roundtable discussed main concerns, risks, and vulnerabilities in relation to data breaches in data spaces and good practices.

The moderators presented a fictional scenario of a health data space where public and private hospitals would share patients' health records for research purposes. According to this scenario, the hospitals share the sensitive personal data in a vulnerable manner from a security perspective.

Several risks and weaknesses were identified by the participants, stemming from a governance perspective but also due to the high complexity (organisational, legal and technological complexity) of data processing in data spaces, which render them attractive targets for attackers.

Even when strong security measures are implemented, massive breaches of personal data may arise in data spaces, with high impacts on the rights and freedoms of data subjects, and also possibly high impact at social level.

Scale is a significant source of risk for data breaches in data Spaces (multiple actors, amount of data) but also complex architectures and data spaces models, where data will be shared across data spaces or in different layers within the same data space. New types of attacks are anticipated. In addition, the participants highlighted the risk for attacks from non-EU adversaries, but also foreign legislation that might enable foreign authorities to request to access the data shared in the data space for example for law enforcement purposes (e.g. US CLOUD Act). That could be the case when implementing services that store personal data in non-EU countries (such international transfers of data) may present risks to the rights and freedoms of data subjects.

Other risks might occur in the process of making the dataset interoperable to a given data space.

The distribution of responsibility among the different actors in data spaces is uncertain. Different architectures and different governance models, in combination with the multitude of actors with different rights and roles will be problematic in a case of a breach, for example caused by a malicious attack. In such a case, there will not be ownership of the problem, and thus the appropriate procedures and measures to report and mitigate the impact of the data breach will not be implemented, at least appropriately.

The legal classification of actors as controllers, joint controllers, processors or sub-processors, will depend on how the governance and infrastructure of the data space is set.

Another issue discussed was the applicability of different legal frameworks in parallel (different incidents notifications requirements). In addition to, and separate from, the notification and communication of personal data breaches under the GDPR, controllers should also be aware of any requirement to notify security incidents under other associated legislation that may apply to them and whether this may also require them to notify the supervisory authority of a personal data breach at the same time.

That will be the case of Digital Operational Resilience Act (DORA) for the financial sector, or the NIS2 Directive that require operators of essential services and digital service providers to notify security incidents to their competent authority. That means that where such incidents are, or become personal data breaches under GDPR (and that is not always the case, as there are security incidents which do not compromise personal data, and vice versa) those operators and providers would be required to notify the Supervisory data protection authority separately from the incident notification requirements of NIS2 and other applicable legal frameworks. Those reporting obligations are following different timeframes, require reporting to different competent, as seen, ant the type of information to be reported also differs.

One key aspect is clear role assignment and responsibilities before the data space is set and any data breach may take place. In the event of a personal data breach, procedures and responsibilities shall be clear in advance, this could be done either with the Terms & Conditions of the dataspace, contractual agreements among the different actors, license agreements, role-based access management, and due diligence obligations.

Next, a risk management framework is crucial. The framework should include a data provenance plan, emergency response plans, but also the actors should engage in different data breach scenarios.

Another recommendation concerns the automation of processes, which, where possible, should detect breaches. Furthermore, a proactive use of Privacy Enhancing Techniques is recommended.

Even with the best cybersecurity standards and security measures in place, personal data breaches will still happen, so PETs (Privacy-Enhancing Technologies) shall be also relevant in data spaces. The term ‘sharing’ shall be understood as “accessing and processing” which is not the same as personal data transfer and copy.

Information and communication to the data subjects is important. Dynamic ways to present information about risks to their rights and freedoms may be a good practice.

The use of widely accepted mechanisms, such as technical standards, certification, and codes of conduct to prove compliance will be useful for data spaces. We should build on what already exists and explore where there are gaps.

Training and awareness of cybersecurity and privacy risks is necessary, before holders, intermediaries, and users access the data spaces. A good practice would be a knowledge transfer platform for dataspace actors.

Finally, data breaches and compliance should be treated as a supply chain problem. The fact that one actor has all necessary technical and organisational measures in place, does not mean this actor will not inherit vulnerabilities or weaknesses from another actor, e.g. a data holder that provided a dataset or an intermediary that curated a dataset. In such a scenario, it must be avoided that responsibility is diluted among the organizations involved in the processing, which must act in a coordinated manner in the management of risks for the rights and freedoms of data subjects.

IV. FINAL CONCLUSIONS

The participants in the working tables shared their views, concerns and ideas with regards to the Data Spaces and the protection of personal data. Attempting to summarize the discussions, it was evident that there are still a number of open questions, main due to the novelty of the concept but also due to the different possibilities for personal data processing and the different actors involved.

The main common element from the discussions was that we need to assess the experience and expertise from applying GDPR principles to existing processing operations and attempt, by analogy, to transfer it to processing operations in data spaces. This might not be as straight forward at the beginning, but it will allow us to take on board all the existing good practises and processes.

Another common element from the discussions was the need for consultation and guidance at National and European level. Data Spaces is a new concept that is still under development while we explore its full potential. During these early deployment phases, stakeholders should be able to share their experiences, practises and identified solutions and be able to consult with regulators.

The last element that was highlighted was the evolution of the technological landscape for data sharing. As new technologies, such as federated learning, are evolving, we need to be able to identify and assess both risks but also opportunities for meeting GDPR principles. In that direction, analyses and good practises would be appreciated but also beneficial.

V. REFERENCES

AEPD-ENISA event “Data Spaces in EU: Synergies between data protection and data spaces, EU challenges and the experiences of Spain” [oct 2023]

AEPD - ENISA conference on Data Spaces